Machine Learning Engineer Nanodegree Capstone Proposal

FORTAS Chihabeddine
March 12th, 2023

Proposal

Domain Background

Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos, and other visual data, and to take action or make recommendations based on that information. If AI allows computers to think, computer vision will allow them to see, observe and understand. [1]

Recently, machine learning in the supply chain industry has proven to be very useful, powerful, and more accurate in various tasks such as inventory management, which can predict demand, and warehouse optimization, which can detect excess and shortages of goods in your store in time. [1]

Amazon Fulfillment Centers are highly active innovation centers that enable Amazon to deliver millions of products to more than 100 countries worldwide using robotics and computer vision technologies.

Distribution centers often use robots to transport items as part of their operations. Items are carried in bins that can hold multiple items. The supply process, from procurement to delivery, is managed by artificial intelligence. Sometimes products are misplaced during processing, resulting in a mismatch between the recorded inventory and the actual items in the bins.

This mismatch can be detected by implementing machine learning technology, which will improve the effectiveness and efficiency of the distribution center.

Problem Statement

Amazon uses a random storage scheme where items are placed into accessible bins with available space, so the contents of each bin are random, rather than organized by specific product types. Thus, each bin image may show only one type of product or a diverse range of products. Occasionally, items are misplaced while being handled, so the contents of some bin images may not match the recorded inventory of that bin.

Now, this project is about building a model that can count the number of objects in each bin. A system like this can be used to track inventory and make sure that delivery consignments have the correct number of items

Datasets and Inputs

Amazon Bin Image Dataset. The dataset contains 500,000 images of bins containing one or more objects. For each image, there is a metadata file containing information about the image like the number of objects, its dimensions, and the type of object. The contents of each bin are random, rather than organized by specific product types. Thus, each bin image may show only one type of product or a diverse range of products. [2]

Images and their associated metadata share simple numerical unique identifiers. For example, the metadata for the image at IMAGE is found at JSON.

Solution Statement

The solution here is to use AWS SageMaker and good machine-learning engineering practices to fetch data from Amazon Bin Image Dataset, preprocess it, and then train a pre-trained model that can classify the image based on the number of objects in the bin. We can count individual instances separately, which means if there are two same objects in the bin, you count them as two. Once we trained the model, we can deploy the model to a SageMaker endpoint and then query it with an image to get a prediction

Benchmark Model

The benchmark for this project is to be able to get a model accuracy of 55.67% or above. This value has been taken from the Amazon Bin Image Dataset Challenge. Following shows the results on validation split. [3]

Accuracy(%)	RMSE(Root Mean Square Error)
55.67	0.930

Quantity	Per class accuracy(%)	Per class RMSE
0	97.7	0.187
1	83.4	0.542
2	67.2	0.710
3	54.9	0.867
4	42.6	1.025
5	44.9	1.311

Evaluation Metrics

For counting task, we going to evaluate our model using two standard metrics, accuracy (precision) and root mean square error (RMSE). 1 is indicator function, and p and g are prediction and ground truth respectively. [3]

Accuracy: $\frac{1}{N}\sum_{i=1}^{N} 1[p_i == g_i]$

RMSE: $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - g_i)^2}$

Project Design

To perform this image classification, we will be using transfer learning to fine-tune a pre-trained CNN model like resnet50 or inception v3 on our dataset. the tasks involved are the following:

1. Setting up AWS Sage Maker: here we setup the environment where we can run our data preparation and training code.
2. Download and preprocess the Amazon Bin Image Dataset.
3. Preparing Data: here we analyze the images to extract meaningful insights, or inspecting for anomalies that could deviate the model training.
4. Hyperparameter Tuning: here we try a range of values for hyperparameter to choose the best parameter for our training.
5. Model Training: now we take best hyperparameter with Multi-Instance Training.
6. Model Deployment: after completing the training we need to deploy our model.
7. Create a lambda function to invoke the endpoint.

Bibliography

[1] O. Kholodenko, "MACHINE LEARNING IN SUPPLY CHAIN," 20 December 2022. [Online]. Available: https://codeit.us/blog/machine-learning-in-supply-chain#:~:text=Machine%20learning%20in%20the%20supply,in%20your%20store%20on%20time..

[2] Amazon, "https://registry.opendata.aws/amazon-bin-imagery/," 2023. [Online]. Available: https://registry.opendata.aws/amazon-bin-imagery/.

[3] silverbottlep, "Amazon Bin Image Dataset(ABID) Challenge," 20 Jully 2017. [Online]. Available: https://github.com/silverbottlep/abid_challenge.

[4] D. DAVE, "Amazon Bin Image Dataset," 2023. [Online]. Available: https://www.kaggle.com/datasets/dhruvildave/amazon-bin-image-dataset.